GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: March 9, 2002, 00:09:21 ; Search time 8498.8 Seconds
(without alignments)
27.817 Million cell updates/sec

Title: US-09-851-670-13
Perfect score: 22
Sequence: 1 cacccgctctctcgacaatgga 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 111874

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_htc:*
10: gb_est1:*
11: gb_est2:*
12: gb_htc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		14.2	64.5	52	10	AW692215	AW692215 NF048H04S
2	c	13.8	62.7	34	13	AQ025206	AQ025206 EP(3)1249
3	c	13	59.1	40	11	W98495	W98495 mg13d12.r1
4		13	59.1	58	10	AA857578	AA857578 of64f12.s
5		12.4	56.4	51	13	AQ074058	AQ074058 EP(2)2050
6	c	12.4	56.4	52	10	AA662933	AA662933 nu92d08.s
7	c	12.4	56.4	52	11	BI175649	BI175649 OSTR051E8
8		12.2	55.5	58	10	AI494282	AI494282 qy98d02.x
9	c	12	54.5	37	10	AI667553	AI667553 fc41g05.x
10	c	12	54.5	50	10	AU102542	AU102542 AU102542
11		12	54.5	50	10	AU103241	AU103241 AU103241
12	c	12	54.5	58	10	AA612377	AA612377 vo03g03.r
13		11.8	53.6	30	13	AZ394609	AZ394609 1M0158F13
14	c	11.8	53.6	37	13	AZ658111	AZ658111 1M0534N12
15		11.8	53.6	43	10	AI795088	AI795088 sb76c01.y
16		11.8	53.6	43	13	AZ834659	AZ834659 2M0117B01
17		11.8	53.6	49	10	AI367448	AI367448 qo76c10.x
18		11.8	53.6	54	13	AZ645918	AZ645918 1M0511L07
19	c	11.8	53.6	58	10	AU008450	AU008450 AU008450
20		11.6	52.7	21	13	AZ976439	AZ976439 2M0251B24
21		11.6	52.7	31	13	TA229E12Q	AL480866 T. brucei
22		11.6	52.7	41	11	R34351	R34351 yh85c03.s1
23		11.6	52.7	44	10	BE378922	BE378922 601237513
24	c	11.6	52.7	50	10	AU103953	AU103953 AU103953
25	c	11.6	52.7	50	10	AU107212	AU107212 AU107212
26		11.6	52.7	50	13	AZ491437	AZ491437 1M0325D04
27		11.6	52.7	55	10	AI935707	AI935707 wo99g08.x
28	c	11.6	52.7	58	10	AA687409	AA687409 ns16a12.s
29		11.4	51.8	27	13	AZ485936	AZ485936 1M0313F07
30		11.4	51.8	32	13	AZ462085	AZ462085 1M0269P08
31		11.4	51.8	33	13	AZ466859	AZ466859 1M0277C24
32	c	11.4	51.8	36	13	BH011404	BH011404 BG01613-5
33	c	11.4	51.8	37	13	AZ616333	AZ616333 1M0446A08
34		11.4	51.8	40	13	AZ463268	AZ463268 1M0272B01
35	c	11.4	51.8	46	10	AA591068	AA591068 vm01e07.r
36		11.4	51.8	50	10	AU102615	AU102615 AU102615
37		11.4	51.8	50	10	AU104278	AU104278 AU104278
38	c	11.4	51.8	50	10	AU104982	AU104982 AU104982
39		11.4	51.8	52	10	AW693240	AW693240 NF062A03S
40	c	11.4	51.8	58	10	AA566959	AA566959 1013 Lobl
41	c	11.4	51.8	59	10	AV834199	AV834199 AV834199
42	c	11.4	51.8	59	10	AW160067	AW160067 SWL4CAK06
43	c	11.4	51.8	59	13	AZ345576	AZ345576 1M0080K16
44		11.4	51.8	60	10	AW693123	AW693123 NF060E03S
45	c	11.4	51.8	60	11	BF019349	BF019349 uy04b12.y

ALIGNMENTS

RESULT 1
AW692215
LOCUS AW692215 52 bp mRNA EST 21-DEC-2000
DEFINITION NF048H04ST1F1000 Developing stem Medicago truncatula cDNA clone NF048H04ST 5', mRNA sequence.
ACCESSION AW692215
VERSION AW692215.2 GI:11957904
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
REFERENCE 1 (bases 1 to 52)
AUTHORS He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell ,C.J., Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Dixon ,R.A.
TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation Medicago truncatula stem library
JOURNAL Unpublished (2000)
COMMENT On Apr 14, 2000 this sequence version replaced gi:7566951.
Contact: Dixon RA
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7302
Fax: 580 221 7380
Email: radixon@noble.org
Insert Length: 909 Std Error: 0.00
Plate: 048 row: H column: 04
Seq primer: TCACACAGGAAACAGCTATGAC.
FEATURES Location/Qualifiers
source 1. .52
/organism="Medicago truncatula"

```
                     /db_xref="taxon:3880"
                     /clone="NF048H04ST"
                     /clone_lib="Developing stem"
                     /tissue_type="stem"
                     /dev_stage="Pooled developmental"
                     /note="Vector: Lambda Zap; Contains a mixture of
                     internodal stem segments"
BASE COUNT       16 a     18 c      1 g     17 t
ORIGIN


  Query Match             64.5%;  Score 14.2;  DB 10;  Length 52;
  Best Local Similarity   84.2%;  Pred. No. 7.7e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 cacccgctctctcgacaat 19
          ||| | ||||||| |||||
Db     34 CACTCTCTCTCTCAACAAT 52


RESULT   2
AQ025206/c
LOCUS       AQ025206       34 bp    DNA             GSS       23-AUG-2000
DEFINITION  EP(3)1249 Drosophila melanogaster EP line Drosophila melanogaster
            genomic Sequence recovered from 5' end of P element, DNA sequence.
ACCESSION   AQ025206
VERSION     AQ025206.1  GI:3265558
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 34)
  AUTHORS   Liao,G.-C., Rehm,E.J. and Rubin,G.M.
  TITLE     Insertion site preferences of the P transposable element in
            Drosophila melanogaster
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3347-3351 (2000)
  MEDLINE   20202638
COMMENT     Contact: Gerald Rubin
            Berkeley Drosophila Genome Project
            University of California, Berkeley
            LSA Building, Berkeley, CA 94720-3200, USA
            Fax: 5106439947
            Email: gerry@fruitfly.berkeley.edu
            Sequence recovery method was inverse PCR.

            Sequence orientation is forward strand relative to 5' end of P
            element

            The P element insertion position is base 27 in the 34 bases. This
            insertion position refers to the first base of the 8 base target
            recognition sequence.
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .34
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone_lib="Drosophila melanogaster EP line"
                     /note="Inverse PCR was performed on Drosophila
                     melanogaster strains each of which contains a single EP
                     transposable element insertion. (The generation of these
                     insertion strains is described in Rorth P, Szabo K, Bailey
                     A, Laverty T, Rehm J, Rubin GM, Weigmann K, Milan M, Benes
                     V, Ansorge W, Cohen SM. 1998. Systematic gain-of-function
                     genetics in Drosophila. Development 6:1049-1057.) The
                     resultant fragment for each strain was directly sequenced
                     to determine the genomic sequence at the site of
                     insertion. Details of the protocols used can be found at
                     http://fruitfly.berkeley.edu/p_disrupt/inverse_pcr.html."
BASE COUNT       4 a      5 c      17 g      8 t
ORIGIN
```

```
  Query Match             62.7%;  Score 13.8;  DB 13;  Length 34;
  Best Local Similarity   88.2%;  Pred. No. 1.1e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 cacccgctctctcgaca 17
          |||||||||||||  ||
Db     26 CACCCGCTCTCTCTGCA 10


RESULT   3
W98495/c
LOCUS       W98495         40 bp    mRNA            EST       16-JUL-1996
DEFINITION  mg13d12.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone IMAGE:423671 5' similar to SW:NUBM_BOVIN P25708
            NADH-UBIQUINONE OXIDOREDUCTASE 51 KD SUBUNIT PRECURSOR ;, mRNA
            sequence.
ACCESSION   W98495
VERSION     W98495.1  GI:1428405
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 40)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:258223
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: ETPrimer
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .40
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:423671"
                     /clone_lib="Soares mouse embryo NbME13.5 14.5"
                     /sex="unknown"
                     /tissue_type="embryo"
                     /dev_stage="13.5-14.5dpc total fetus"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTT
                     T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                     14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                     State Univ., from 2 ]; double-stranded cDNA was ligated to
                     Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT7T3 vector. Library went through one round of
                     normalization, and was constructed by Bento Soares and
                     M.Fatima Bonaldo.  "
BASE COUNT       8 a      12 c      12 g      8 t
ORIGIN
```

```
  Query Match            59.1%;  Score 13;  DB 11;  Length 40;
  Best Local Similarity  76.2%;  Pred. No. 2.6e+04;
  Matches   16;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 cacccgctctctcgacaatgg 21
          || || || |||| ||| |||
Db     32 CAGCCACTGTCTCCACATTGG 12


RESULT   4
AA857578
LOCUS       AA857578       58 bp    mRNA            EST       21-APR-1998
DEFINITION  of64f12.s1 NCI_CGAP_Co8 Homo sapiens cDNA clone IMAGE:1435151 3',
            mRNA sequence.
ACCESSION   AA857578
VERSION     AA857578.1  GI:2945880
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1149   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 56.
FEATURES             Location/Qualifiers
     source          1. .58
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1435151"
                     /clone_lib="NCI_CGAP_Co8"
                     /tissue_type="adenocarcinoma"
                     /lab_host="DH10B"
                     /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; 1st strand cDNA was prepared from
                     colon adenocarcinoma, and was then primed with a Not I -
                     oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                     RI adaptors (Pharmacia), digested with Not I and cloned
                     into the Not I and Eco RI sites of the modified pT7T3
                     vector.  Library is normalized. Library was constructed by
                     Bento Soares and M. Fatima Bonaldo. "
BASE COUNT       16 a     14 c      7 g     21 t
ORIGIN


  Query Match            59.1%;  Score 13;  DB 10;  Length 58;
  Best Local Similarity  76.2%;  Pred. No. 2.9e+04;
  Matches   16;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 cacccgctctctcgacaatgg 21
          ||||  |||||  || |||||
Db     31 CACCACCTCTCATGAAAATGG 51


RESULT   5
AQ074058
LOCUS       AQ074058       51 bp    DNA             GSS       23-AUG-2000
DEFINITION  EP(2)2050-5prime Drosophila melanogaster EP line Drosophila
            melanogaster genomic Sequence recovered from 5' end of P element,
            DNA sequence.
ACCESSION   AQ074058
VERSION     AQ074058.1  GI:3403308
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Liao,G.-C., Rehm,E.J. and Rubin,G.M.
  TITLE     Insertion site preferences of the P transposable element in
            Drosophila melanogaster
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3347-3351 (2000)
  MEDLINE   20202638
COMMENT     Contact: Gerald Rubin
            Berkeley Drosophila Genome Project
            University of California, Berkeley
            LSA Building, Berkeley, CA 94720-3200, USA
            Fax: 5106439947
            Email: gerry@fruitfly.berkeley.edu
            Sequence recovery method was inverse PCR.

            Sequence orientation is forward strand relative to 5' end of P
            element

            The P element insertion position is base 44 in the 51 bases. This
            insertion position refers to the first base of the 8 base target
            recognition sequence.
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone_lib="Drosophila melanogaster EP line"
                     /note="Inverse PCR was performed on Drosophila
                     melanogaster strains each of which contains a single EP
                     transposable element insertion. (The generation of these
                     insertion strains is described in Rorth P, Szabo K, Bailey
                     A, Laverty T, Rehm J, Rubin GM, Weigmann K, Milan M, Benes
                     V, Ansorge W, Cohen SM. 1998. Systematic gain-of-function
                     genetics in Drosophila. Development 6:1049-1057.) The
                     resultant fragment for each strain was directly sequenced
                     to determine the genomic sequence at the site of
                     insertion. Details of the protocols used can be found at
                     http://fruitfly.berkeley.edu/p_disrupt/inverse_pcr.html."
BASE COUNT       10 a     18 c      7 g     16 t
ORIGIN


  Query Match            56.4%;  Score 12.4;  DB 13;  Length 51;
  Best Local Similarity  72.7%;  Pred. No. 5.3e+04;
  Matches   16;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 cacccgctctctcgacaatgga 22
          | || |||||||  |||| | |
Db      2 CGCCGGCTCTCTTTACAACGCA 23


RESULT   6
AA662933/c
LOCUS       AA662933       52 bp    mRNA            EST       12-NOV-1997
DEFINITION  nu92d08.s1 NCI_CGAP_Pr22 Homo sapiens cDNA clone IMAGE:1218159 3'
            similar to TR:G929921 G929921 RNA POLYMERASE II SUBUNIT HSRPB7. ;,
            mRNA sequence.
ACCESSION   AA662933
VERSION     AA662933.1  GI:2616924
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 52)
```

```
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .52
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1218159"
                     /clone_lib="NCI_CGAP_Pr22"
                     /sex="male"
                     /tissue_type="normal prostate"
                     /lab_host="DH10B"
                     /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; 1st strand cDNA was prepared
                     from normal prostate bulk tissue, and was then primed with
                     a Not I - oligo(dT) primer. Double-stranded cDNA was
                     ligated to Eco RI adaptors (Pharmacia), digested with Not
                     I and cloned into the Not I  and Eco RI sites of the
                     modified pT7T3 vector.  Library is normalized, and was
                     constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT       12 a      14 c      14 g      12 t
ORIGIN


  Query Match             56.4%;  Score 12.4;  DB 10;  Length 52;
  Best Local Similarity   72.7%;  Pred. No. 5.4e+04;
  Matches   16;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 cacccgctctctcgacaatgga 22
           | |||  | ||  |||||||||
Db      39 CCCCCTGTTTCAAGACAATGGA 18


RESULT   7
BI175649/c
LOCUS       BI175649       52 bp    mRNA            EST       09-JUL-2001
DEFINITION  OSTR051E8_1 AD-wrmcDNA Caenorhabditis elegans cDNA similar to
            F42H10.4, mRNA sequence.
ACCESSION   BI175649
VERSION     BI175649.1  GI:14641452
KEYWORDS    EST.
SOURCE      Caenorhabditis elegans.
  ORGANISM  Caenorhabditis elegans
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
            ; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1  (bases 1 to 52)
  AUTHORS   Reboul,J., Vaglio,P., Tzellas,N., Thierry-Mieg,N., Moore,T.,
            Jackson,C., Shin-i,T., Kohara,Y., Thierry-Mieg,D., Thierry-Mieg,J.,
            Lee,H., Hitti,J., Doucette-Stamm,L., Hartley,J.L., Temple,G.F.,
            Brasch,M.A., Vandenhaute,J., Lamesch,P.E., Hill,D.E. and Vidal,M.
  TITLE     Open-reading-frame sequence tags (OSTs) support the existence of at
            least 17,300 genes in C. elegans
  JOURNAL   Nat. Genet. 27 (3), 332-336 (2001)
  MEDLINE   21135099
COMMENT     Contact: Reboul J, Vaglio P
            Marc Vidal Laboratory
            Dana Farber Cancer Institute
            44 Binney Street, Boston, MA 02115, USA
            Tel: 617 632 5180
            Fax: 617 632 2425
            Email: Jerome_Reboul@dfci.harvard.edu
            Sequence tag of Gateway entry clones. The primers used were
            designed on the predicted protein encoding ORF. C. elegans ORFeome
            cloning project : Contact jerome_reboul@dfci.harvard.edu or
            philippe_vaglio@dfci.harvard.edu
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .52
                     /organism="Caenorhabditis elegans"
                     /strain="N2"
                     /db_xref="taxon:6239"
                     /clone_lib="AD-wrmcDNA"
                     /sex="Hermaphrodite and male"
                     /tissue_type="whole animal"
                     /dev_stage="mixed stage"
                     /note="The AD-wrmcDNA library was generated with poly(A)+
                     RNA isolated from both hermaphrodite and male N2 worms of
                     all larval stages, embryos, adults and dauers and the
                     subsequent generation of cDNAs by poly(A) priming. The
                     cDNAs were cloned into pPC86"
BASE COUNT       17 a      15 c       7 g      13 t
ORIGIN


  Query Match             56.4%;  Score 12.4;  DB 11;  Length 52;
  Best Local Similarity   72.7%;  Pred. No. 5.4e+04;
  Matches   16;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 cacccgctctctcgacaatgga 22
           || | | || || || ||||||
Db      40 CAACTGGTCACTGGATAATGGA 19


RESULT   8
AI494282
LOCUS       AI494282       58 bp    mRNA            EST       17-MAR-1999
DEFINITION  qy98d02.x1 NCI_CGAP_Brn25 Homo sapiens cDNA clone IMAGE:2020035 3'
            similar to TR:Q34192 Q34192 NADH DEHYDROGENASE SUBUNIT 5. ;, mRNA
            sequence.
ACCESSION   AI494282
VERSION     AI494282.1  GI:4395285
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute / National Institute of Neurological
            Disorders and Stroke, Brain Tumor Genome Anatomy Project
            (CGAP/BTGAP), Tumor Gene Index
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
            Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1046    Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .58
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2020035"
```

```
                     /clone_lib="NCI_CGAP_Brn25"
                     /tissue_type="anaplastic oligodendroglioma"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATAGGTTTTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library is normalized, and was constructed by Bento
                     Soares and M.Fatima Bonaldo."
BASE COUNT        9 a      21 c       1 g      27 t
ORIGIN


  Query Match             55.5%;  Score 12.2;  DB 10;  Length 58;
  Best Local Similarity   82.4%;  Pred. No. 6.8e+04;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 cccgctctctcgacaat 19
           ||| ||||||| || ||
Db      34 CCCTCTCTCTCTACTAT 50


RESULT   9
AI667553/c
LOCUS       AI667553       37 bp    mRNA            EST       07-JUN-2001
DEFINITION  fc41g05.x1 Zebrafish WashU MPIMG EST Danio rerio cDNA clone
            IMAGE:3723992 3' similar to SW:PAB1_MOUSE P29341
            POLYADENYLATE-BINDING PROTEIN 1 ;, mRNA sequence.
ACCESSION   AI667553
VERSION     AI667553.1  GI:4805909
KEYWORDS    EST.
SOURCE      zebrafish.
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Rasborinae; Danio.
REFERENCE   1  (bases 1 to 37)
  AUTHORS   Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
            ,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
            ,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
            Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
            Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
            and Wilson,R.
  TITLE     WashU Zebrafish EST Project 1998
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Stephen L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: zbrafish@watson.wustl.edu
            cDNA Library Preparation: Matthew Clark. cDNA Library Arrayed by:
            Matthew Clark. DNA Sequencing by: Washington University Genome
            Sequencing Center Clone distribution: Genome Systems, St. Louis,
            Missouri (web address: www.genomesystems.com) (email contact:
            info@genomesystems.com) and Research Genetics, Huntsville, Alabama
            (web address: www.resgen.com) (email contact: info@resgen.com) and
            RessourcenZentrumPrimarDatenbank, Berlin, Germany (web address:
            www.rzpd.de)
            Possible reversed clone: similarity on wrong strand
            Seq primer: T7 ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .37
                     /organism="Danio rerio"
                     /db_xref="taxon:7955"
                     /clone="IMAGE:3723992"
                     /clone_lib="Zebrafish WashU MPIMG EST"
                     /sex="mixed"
                     /tissue_type="26 somite embryos, adult livers, shield
                     stage embryos"
                     /lab_host="XL1-blue MRF"
                     /note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI; 1st
                     strand cDNA was primed with a Not I - oligo(dT)15 primer
                     [5'pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT3']
                     double-stranded cDNA was ligated to Sal I adaptors (BRL),
                     digested with Not I and cloned into the Not I and Sal I
                     sites of the pSPORT1 vector (BRL). Library was constructed
                     by Matthew Clark (Lehrach lab; ICRF, London and Max Planck
                     Institut fuer Molekulare Genetik,Berlin). cDNAs for EST
                     analysis were selected following oligonucleotide
                     hybridization fingerprinting of arrayed clones from
                     zebrafish late somitogenesis (26 ss), adult liver or
                     embryonic shield stage (5.6 h) libraries. Fingerprint
                     data were used to computationally cluster cDNAs, and a
                     single cDNA from each cluster was chosen for sequencing.
                     In some cases multiple members of the same cluster were
                     sequenced to assess clustering parameters or single clones
                     were sequenced additional times to assess quality
                     control."
BASE COUNT        8 a      9 c      13 g      7 t
ORIGIN


  Query Match             54.5%;  Score 12;  DB 10;  Length 37;
  Best Local Similarity   75.0%;  Pred. No. 7.6e+04;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 cccgctctctcgacaatgga 22
           |||||||||  ||  ||| |
Db      37 CCCGCTCTCCAGAACATGCA 18


RESULT  10
AU102542/c
LOCUS       AU102542       50 bp    mRNA            EST       05-APR-2001
DEFINITION  AU102542 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HRC08226, mRNA sequence.
ACCESSION   AU102542
VERSION     AU102542.1  GI:13552063
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
            ,K., Suyama,A. and Sugano,S.
  TITLE     Fine Structural analysis of transcription start sites of human
            mRNAs using full-length enriched and 5'-end enriched cDNA libraries
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HRC08226"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT        7 a      12 c      20 g      11 t
ORIGIN


  Query Match             54.5%;  Score 12;  DB 10;  Length 50;
  Best Local Similarity   75.0%;  Pred. No. 8.2e+04;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;
```

```
Qy       1 cacccgctctctcgacaatg 20
           || ||||  || | |||||||
Db      24 CAGCCGCCATCCCAACAATG 5


RESULT  11
AU103241
LOCUS       AU103241       50 bp    mRNA            EST       05-APR-2001
DEFINITION  AU103241 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HSI07361, mRNA sequence.
ACCESSION   AU103241
VERSION     AU103241.1  GI:13552762
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
            ,K., Suyama,A. and Sugano,S.
  TITLE     Fine Structural analysis of transcription start sites of human
            mRNAs using full-length enriched and 5'-end enriched cDNA libraries
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HSI07361"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT       12 a     17 c      8 g     13 t
ORIGIN


  Query Match             54.5%;  Score 12;  DB 10;  Length 50;
  Best Local Similarity   75.0%;  Pred. No. 8.2e+04;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 cacccgctctctcgacaatg 20
           ||| | |||||  | |||||
Db       6 CACTCTCTCTCCTGTCAATG 25


RESULT  12
AA612377/c
LOCUS       AA612377       58 bp    mRNA            EST       01-OCT-1997
DEFINITION  vo03g03.r1 Stratagene mouse skin (#937313) Mus musculus cDNA clone
            IMAGE:1040404 5', mRNA sequence.
ACCESSION   AA612377
VERSION     AA612377.1  GI:2461055
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:580428
            Seq primer: -28ml3 rev1 ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .58
                     /organism="Mus musculus"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1040404"
                     /clone_lib="Stratagene mouse skin (#937313)"
                     /sex="females"
                     /tissue_type="whole skin"
                     /dev_stage="11 weeks old"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: skin; Vector: pBluescript SK-; Site_1: EcoRI
                     ; Site_2: XhoI; Cloned unidirectionally.  Primer: Oligo
                     dT. Whole skin from 11 week old C57BL/6 female mice.
                     Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5'
                     adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor
                     sequence: 5' CTCGAGTTTTTTTTTTTTTTTTT 3'"
BASE COUNT       21 a     15 c      8 g     14 t
ORIGIN


  Query Match             54.5%;  Score 12;  DB 10;  Length 58;
  Best Local Similarity   75.0%;  Pred. No. 8.5e+04;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 cccgctctctcgacaatgga 22
           |||||| | |  | ||||||
Db      29 CCCGCTGTATTAAAAATGGA 10


RESULT  13
AZ394609
LOCUS       AZ394609       30 bp    DNA             GSS       03-OCT-2000
DEFINITION  1M0158F13F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0158F13 F, DNA sequence.
ACCESSION   AZ394609
VERSION     AZ394609.1  GI:10509681
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0158  row: F  column: 13
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 30.
```

```
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0158F13"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT        3 a      8 c      9 g      10 t
ORIGIN


  Query Match             53.6%;  Score 11.8;  DB 13;  Length 30;
  Best Local Similarity   86.7%;  Pred. No. 9e+04;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 cacccgctctctcga 15
           |||||| ||||| ||
Db       7 CACCCGGTCTCTTGA 21


RESULT  14
AZ658111/c
LOCUS       AZ658111       37 bp    DNA             GSS       14-DEC-2000
DEFINITION  1M0534N12R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0534N12 R, DNA sequence.
ACCESSION   AZ658111
VERSION     AZ658111.1  GI:11795257
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 37)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0534  row: N  column: 12
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 37.
FEATURES             Location/Qualifiers
     source          1. .37
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0534N12"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       13 a      1 c      19 g      4 t
ORIGIN


  Query Match             53.6%;  Score 11.8;  DB 13;  Length 37;
  Best Local Similarity   86.7%;  Pred. No. 9.4e+04;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 cccgctctctcgaca 17
           ||| ||||||| |||
Db      31 CCCCCTCTCTCCACA 17


RESULT  15
AI795088
LOCUS       AI795088       43 bp    mRNA            EST       13-DEC-1999
DEFINITION  sb76c01.y1 Gm-c1010 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
            Gm-c1010-889 5' similar to SW:ITRA_SOYBN P01070 TRYPSIN INHIBITORS
            A AND C PRECURSOR ;, mRNA sequence.
ACCESSION   AI795088
VERSION     AI795088.1  GI:5342804
KEYWORDS    EST.
SOURCE      soybean.
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 43)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
            ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
            Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
            ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
            ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
            ,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
```

This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
High quality sequence stop: 1.

```
FEATURES             Location/Qualifiers
     source          1. .43
                     /organism="Glycine max"
                     /db_xref="taxon:3847"
                     /clone="GENOME SYSTEMS CLONE ID: Gm-c1010-889"
                     /clone_lib="Gm-c1010"
                     /tissue_type="young cotyledons of greenhouse grown plants"
                     /dev_stage="2cm long 12 week old"
                     /lab_host="XL10-Gold"
                     /note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
                     XhoI; This cDNA library was constructed from mRNA isolated
                     from immature cotyledons (100-200mgs) of old greenhouse
                     grown plants. The cDNA library was prepared using the
                     Stratagene pBluescript II SK(+) library construction kit.
                     Complementary DNA was synthesized from mRNA using a primer
                     consisting of a poly (dT) sequence with a XhoI restriction
                     site. EcoRI adapters were ligated to the blunt-ended cDNA
                     fragments followed by XhoI digestion. The cDNA fragments
                     were directionally cloned into the EcoRI-XhoI restriction
                     site of the pBluescript vector. The ligated cDNA fragments
                     were transformed into XL10-Gold host cells. This library
                     was constructed by Dr. Lila Vodkin and Dr. Anu Khanna."
BASE COUNT       11 a     10 c      7 g     15 t
ORIGIN


  Query Match             53.6%;  Score 11.8;  DB 10;  Length 43;
  Best Local Similarity   86.7%;  Pred. No. 9.8e+04;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       7 ctctctcgacaatgg 21
           | ||||||| |||||
Db       7 CCCTCTCGAAAATGG 21


Search completed: March  9, 2002, 00:09:23
Job time: 11039 sec
```